Proceedings of the American Academy of Arts and Sciences.

Vol. 66. No. 6.—March, 1931.

THE PROBLEMS OF LAGRANGE AND MAYER WITH VARIABLE END POINTS.

By Marston Morse and Sumner Byron Myers.

THE PROBLEMS OF LAGRANGE AND MAYER WITH VARIABLE END POINTS.

By Marston Morse and Sumner Byron Myers.

1. **Introduction.** In developing sufficient conditions[1] in this problem, Morse found it convenient to give the finite conditions on the end points of the curves by representing these end points as functions of parameters rather than by expressing these conditions in the usual implicit form. This eliminated much of the algebra involved in the second variation and brought to the fore the essentials of the problem.

From the point of view of simplicity and elegance this was compatible only with a treatment of the Euler equations and the transversality conditions in corresponding form.

This was the first objective of the present paper. Incidentally the transversality condition is given a form that has apparently never appeared in the general problem, a form that reduces to the classic condition in the ordinary problem.

The paper also contains a new criterion for "normalcy" for the case of variable end points. This criterion has the advantage that it is near to the criterion given in the fixed end point problem, and has an immediate geometric interpretation which leads to various new theorems.

The form of the problem is an extension of the one used by Bolza.[2] The expression to be minimized is the sum of an integral and a function of a set of parameters which determine the end points, but which are not themselves in general uniquely determined by the end points. It is more general than any form hitherto treated at length, unless one resorts to devices such as introducing new differential conditions or terminal conditions. Such devices often change normal extremals to anormal extremals, or vice versa, and since sufficient conditions usually require normalcy, such changes lead to new difficulties. Furthermore, the generality of the form of the problem used here corresponds to an actual need in the treatment of the second variation.

The authors have made extensive use of the lectures of Bliss.[3]

2. **The problem.** In the space of the variables x and

$$(y) = (y_1, \cdots, y_n)$$

let there be given a curve E

$$y_i = \bar{y}_i(x) \qquad a^1 \leqq x \leqq a^2 \qquad (i = 1, \cdots, n) \tag{2.1}$$

of class D'*. Points neighboring the initial and final end points of E will be denoted respectively by

$$(x^s, y_1{}^s, \cdots, y_n{}^s) = (x^s, y^s), \tag{2.2}$$

where $s = 2$ at the final end point and $s = 1$ at the initial end point.

We consider curves of class D' neighboring E. Such curves will be called *differentially admissible* if they satisfy m differential equations of the form

$$\phi_\beta(x, y, y') = 0 \qquad (\beta = 1, \cdots, m). \qquad m < n\dagger \tag{2.3}$$

We suppose that E is differentially admissible and that along E the functional matrix of the functions (2.3) with respect to the variables y_i' is of rank m.

A curve neighboring E will be said to be *terminally admissible* if its end points are given for some value of (α) by the functions

$$x^s = x^s(\alpha_1, \cdots, \alpha_r) \qquad y_i{}^s = y_i{}^s(\alpha_1, \cdots, \alpha_r), \tag{2.4}$$

where these functions of (α) are defined for (α) near (0), and reduce to the end points of E for $(\alpha) = (0)$.

A curve that is both differentially and terminally admissible will be called *admissible*. A set (α) neighboring $(\alpha) = (0)$ will be called *admissible* if it determines through (2.4) the end points of some admissible curve. We note that infinitely many sets (α) might determine the same admissible curve.

We seek conditions under which E and the set $(\alpha) = (0)$ afford a minimum for the expression

$$J = \int_{x^1}^{x^2} f(x, y, y')dx + \theta(\alpha) \tag{2.5}$$

among admissible sets (α) and corresponding admissible curves of class D'. The functions f and ϕ_β are to be of class C'' while the function $\theta(\alpha)$ and the end point functions in (2.4) need be of no more than class C'.

We note that $\theta(\alpha)$ *may contain variables which do not explicitly occur in the terminal conditions* (2.4).

---

* See Bolza, "Vorlesungen über Variationsrechnung," p. 63.

† Here m could equal n as well, and the following developments would hold. One would not have to "extend" the differential equations $\phi_\beta = 0$ nor prove any of the "multipliers" zero.

3. **The variations.** Suppose that we have a family of differentially admissible curves

$$y_i = y_i(x, t) \qquad x_0^1(t) \leqq x \leqq x_0^2(t) \quad (i = 1, \cdots, n) \tag{3.1}$$

containing E for $t = 0$. The *variations* of the family (3.1) along E with respect to t are defined to be the following derivatives:

$$\gamma^s = x_{0t}{}^s(0) \qquad \eta_{i}(x) = y_{it}(x, 0) \tag{3.2}$$

where the subscript t denotes differentiation with respect to t. These variations satisfy the *differential equations of variation**

$$\Phi_\beta \equiv \phi_{\beta y_i}\eta_i + \phi_{\beta y_i'}\eta_i' = 0 \tag{3.3}$$

obtained by differentiating (2.3) with respect to t and setting $t = 0$. Here (x, y, y') are to be taken on E.

A set (α) in (2.4) determines a set of end points. Let us take a set of points (α) of the form

$$\alpha_h = \alpha_h(t) \qquad \alpha_h(0) = 0, \qquad (h = 1, \cdots, r) \tag{3.4}$$

where the functions $\alpha_h(t)$ are of class C' for t near 0. Suppose that the family (3.1) takes on the end points determined by (α) in (3.4). That is, we suppose we have, subject to (3.4), the following identities in t:

$$\begin{aligned} x_0{}^s(t) &= x^s(\alpha) \\ y_i[x_0{}^s(t), t] &= y_i{}^s(\alpha). \end{aligned} \tag{3.5}$$

The curves of the family (3.1) are now also terminally admissible, and we include among the variations the following derivatives:

$$u_h = \alpha_h{}'(0).$$

These, together with the variations (3.2) satisfy what we shall call the *terminal equations of variation*

$$\begin{aligned} \gamma^s &= x_h{}^s u_h \quad (s = 1, 2;\ h = 1, \cdots, r;\ i = 1, \cdots, n) \\ \eta_i{}^s &= (y_{ih}{}^s - \overline{y}_i{}'{}^s x_h{}^s)u_h \qquad\qquad u_h = \alpha_h{}'(0) \end{aligned} \tag{3.6}$$

obtained by differentiation of (3.5) with respect to t and setting $t = 0$. In (3.6) the superscript s means evaluation at the corresponding end of E and the subscript h attached to x^s or $y_i{}^s$ means differentiation with respect to α_h.

* The summation convention of tensor analysis is used throughout.

A set γ^s, $\eta_i(x)$, where the set γ^s consists of two arbitrary constants and the functions $\eta_i(x)$ are n functions of class D' which satisfy (3.3), will be called a set of *differentially admissible variations.*

A set of numbers γ^s, $\eta_i{}^s$ which satisfy (3.6) for some set $(u) = (u_1 \cdots, u_r)$ will be called a set of *terminally admissible variations,* and will be said to be determined by the set (u).

With these definitions, we can now prove the following theorem:

THEOREM 1. *For every set of differentially admissible variations* γ^s, $\eta_i(x)$ *there exists a one-parameter family*

$$y_i = y_i(x, t) \qquad x_0{}^1(t) \leqq x \leqq x_0{}^2(t) \tag{3.7}$$

of differentially admissible arcs containing E *for* $t = 0$ *and having* γ^s, $\eta_i(x)$ *as its variations along* E. *The functions* $y_i(x, t)$ *are continuous and have continuous derivatives with respect to* t *for all values* (x, t) *near those defining* E, *and the derivatives* $y_{ix}(x, t)$ *have the same properties except possibly at the values of* x *defining corners of* E, *while the functions* $x_0{}^s(t)$ *are of class* C' *for* t *near* 0.

This theorem and its proof are due to Bliss.* For purposes of reference we will reproduce the essentials of the proof here.

First we enlarge the system $\phi_\beta = 0$ to the extended system of differential equations

$$\phi_\beta = 0 \qquad \phi_\tau = z_\tau(x) \quad (\beta = 1, \cdots, m; \tau = m+1, \cdots, n) \tag{3.8}$$

where ϕ_τ are new functions of (x, y, y') such that the functional determinant of the extended system with respect to the variables y_i' is different from 0 along E† and the functions $z_\tau(x)$ are defined by (3.8) upon substituting in ϕ_τ the functions $\overline{y}_i(x)$ defining E.

From this extended system we obtain an extended system of differential equations of variation

$$\Phi_\beta = 0 \qquad \Phi_\tau = \zeta_\tau(x) \quad (\beta = 1, \cdots, m; \tau = m+1, \cdots, n) \tag{3.9}$$

along E where the functions $\zeta_\tau(x)$ are defined by these equations corresponding to a set of functions $\eta_i(x)$ of any set of differentially admissible variations.

---

* See Bliss,[3] "The Problem of Lagrange in the Calculus of Variations," loc. cit., p. 4–6, and p. 20.

† For a proof of the possibility of this enlargement see Bliss, "The Problem of Mayer with Variable End Points," Transactions of the American Mathematical Society, vol. 19 (1918), p. 312.

Let the set γ^s, $\eta_i(x)$ be a set of differentially admissible variations for E, and $\zeta_\tau(x)$ the corresponding set defined by (3.9). The existence theorems for differential equations applied to the system

$$\phi_\beta = 0 \qquad \phi_\tau = z_\tau(x) + t\zeta_\tau(x) \tag{3.10}$$

furnish a one-parameter family of continuous solutions

$$y_i = y_i(x, t) \tag{3.11}$$

where

$$x_0^1(t) \leqq x \leqq x_0^2(t) \qquad x_0^s(t) = a^s + t\gamma^s$$

and where the initial values are given by

$$y_i(a^1, t) = \bar{y}_i(a^1) + t\eta_i(a^1)$$

The solutions are required to be continuous at the corner values of x. This family contains the arc E for $t = 0$, and has the prescribed variations γ^s. Its η variations have the initial values $\eta_i(a^1)$ and satisfy (3.9) with the functions $\zeta_\tau(x)$, and hence coincide with the prescribed functions $\eta_{i\tau}(x)$.

In a similar manner the following corollary can be proved:

For a matrix

$$\begin{vmatrix} \gamma_1{}^1 \cdots \gamma_q{}^1 \\ \gamma_1{}^2 \cdots \gamma_q{}^2 \\ \eta_{11} \cdots \eta_{1q} \\ \cdots\cdots\cdots \\ \eta_{n1} \cdots \eta_{nq} \end{vmatrix}$$

where each column is a set of differentially admissible variations, there exists a q-parameter family of differentially admissible arcs

$$y_i = y_i(x, e_1, \cdots, e_q) \quad x_0^1(e_1, \cdots, e_q) \leqq x \leqq x_0^2(e_1, \cdots, e_q) \tag{3.12}$$

containing the arc E for $(e_1, \cdots, e_q) = (0, \cdots 0)$ *and having* $\gamma_p{}^s$, η_{ip} *as its variations along E with respect to* e_p. *The continuity properties are similar to those of the preceding theorem.*

4. **The Euler equations and the transversality condition.** Let us substitute in the integral in J the functions (3.7) defining a one-parameter family of differentially admissible arcs containing E for $t = 0$. If we take the parameters (α) thus,

$$\alpha_h = tu_h$$

where u_h is a set of r arbitrary constants, and substitute them in $\theta(\alpha)$, then J becomes a function of t defined by the formula

$$J(t) = \int_{x^1(t)}^{x^2(t)} f[x, y(x, t), y'(x, t)]dx + \theta(tu). \tag{4.1}$$

The derivative of this function with respect to t at the value $t = 0$ has the value*

$$V(\gamma, \eta, u) = \int_{a^1}^{a^2} (f_{y^i}\eta_i + f_{y_i'}\eta_i')dx + (f^s\gamma^s)_1^2 + \theta_h u_h \tag{4.2}$$

where f^1 and f^2 are the values of f at the respective end points of E, and the arguments of the derivatives are those defining E.

Let us obtain a more useful form of $V(\gamma, \eta, u)$. Let F be defined by the equation

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1\phi_1 + \cdots + \lambda_n\phi_n$$

where λ_0 is a constant and $\lambda_i(x)$ are functions of x in the interval $(x^1\ x^2)$. The value of $\lambda_0 V(\gamma, \eta, u)$ is not changed if we add the sum

$$\lambda_\beta\Phi_\beta + \lambda_\tau(\Phi_\tau - \zeta_\tau)$$

to its integrand, since the variations $\eta_i(x)$ satisfy (3.9). We obtain the equation

$$\lambda_0 V(\gamma, \eta, u) = \int_{a^1}^{a^2} (F_{y_i}\eta_i + F_{y_i'}\eta_i' - \lambda_\tau\zeta_\tau)dx + (\lambda_0 f^s\gamma^s)_1^2 + \lambda_0\theta_h u_h. \tag{4.3}$$

So far the functions $\lambda_i(x)$ have been arbitrary. With the admissible curve E we now associate a constant λ_0 and multipliers $\lambda_i(x)$ so that the equations

$$F_{y_i'} = \int_{a^1}^{x} F_{y_i}dx + c_i \qquad (i = 1, \cdots, n) \tag{4.4}$$

are satisfied for an arbitrarily selected set of constants λ_0, c_i and (x, y, y') on E. This is possible since if we introduce the new variables.

$$v_i = F_{y_i'} = \lambda_0 f_{y_i'} + \lambda_1\phi_{1y_i'} + \cdots + \lambda_n\phi_{ny_i'} \quad (i = 1, \cdots, n) \tag{4.4a}$$

* Here $(\ )_1^2$ means the value of $(\ )$ for $s = 2$, minus its value for $s = 1$.

equations (4.4) are equivalent to the following equations:

$$\frac{dv_i}{dx} = F_{v_i} = A_{i1}v_1 + \cdots + A_{in}v_n + B_i \quad (i = 1, \cdots, n) \tag{4.4b}$$
$$v_i(a^1) = c_i$$

The coefficients A_{ij}, B_i are determined by solving (4.4a) for $\lambda_1, \cdots, \lambda_n$ and substituting in F_{v_i}, and they are continuous except possibly at the values of x defining corners of E. Hence equations (4.4b) have unique continuous solutions $v_i(x)$ which are of class C' except possibly at the corner values of x. Equations (4.4a) then determine the functions $\lambda_i(x)$, continuous except possibly at the corner values of x.

By means of equations (4.3), after integrating by parts we find that

$$\lambda_0 V(\gamma, \eta, u) = -\int_{a^1}^{a^2} \lambda_r \eta_r dx + (\lambda_0 f^s \gamma^s + \eta_i{}^s F^s{}_{y_i'})_1^2 + \lambda_0 \theta_h u_h. \tag{4.5}$$

This equation holds for arbitrary values of λ_0 and $F_{y_i'}{}^1$, for any set of differentially admissible variations $\gamma^s, \eta_i(x)$, and for any set of numbers u_h.

Consider the family (3.12) of the corollary of § 3, letting $q = 2n + 3$. Let

$$u_{hj} \quad (h = 1, \cdots, r; j = 1, \cdots, 2n+3)$$

be $2n + 3$ sets of arbitrary constants. Take the parameters α_h of (2.4) as follows:

$$\alpha_h = e_j u_{hj} \quad (h = 1, \cdots, r; j = 1, \cdots, 2n+3). \tag{4.6}$$

If we substitute the functions defining the family (3.12) in the integral in J, and the parameters (4.6) in $\theta(\alpha)$, then J becomes a function $J(e_1, \cdots, e_{2n+3})$, which when $(e_1, \cdots, e_{2n+3}) = (0, \cdots, 0)$ takes on the value J_0, its value along E. The $2n + 3$ equations

$$J(e_1, \cdots, e_{2n+3}) - J_0 = w$$
$$x_0{}^s(e_1, \cdots, e_{2n+3}) - x^s(e_j u_{1j}, \cdots, e_j u_{rj}) = 0 \tag{4.7}$$
$$y_i[x_0{}^s(e_1, \cdots, e_{2n+3}), e_1, \cdots, e_{2n+3}] - y_i{}^s(e_j u_{1j}, \cdots, e_j u_{rj}) = 0$$

all except the first of which are conditions for terminal admissibility, have the initial solution

$$(w, e_1, \cdots, e_{2n+3}) = (0, 0, \cdots, 0).$$

If the jacobian of the left-hand sides of (4.7) with respect to $e_1, \cdots, e_{2n+3}$ were different from zero at this solution, then the implicit function theorems show that (4.7) would have solutions $(e_1, \cdots, e_{2n+3})$ neighboring $(0, \cdots, 0)$ for all values of w in the neighborhood of $w = 0$. In that case, for negative values of w there would be admissible arcs in the family (3.12) giving J a value smaller than J_0, which is impossible if E is a minimizing arc.

Hence, if E is a minimizing arc, we conclude that the jacobian must be zero.

Let us set

$$\frac{\partial J}{\partial e_j} = V_j \qquad (e_j = 0; j = 1, \cdots, 2n+3)$$

where V_j is $V(\gamma, \tau_i, u)$ evaluated for the variations (γ, τ_i) arising from the family (3.12) for variation of e_j, and (u) stands for the set u_{hj}, $(h = 1, \cdots, r)$. Then the jacobian in question has as the jth of its $2n + 3$ columns:

$$\begin{array}{lll} V_j & 1 \text{ row} & \\ \gamma_j{}^s - x_h{}^s u_{hj} & 2 \text{ rows} & (s = 1, 2) \\ \overline{y}_i{}'^s \gamma_j{}^s + \tau^s{}_{ij} - y^s{}_{ih} u_{hj} & 2n \text{ rows} & (s = 1, 2; i = 1, \cdots, n). \end{array} \tag{4.8}$$

Let us use the notation $\overline{\gamma}_j{}^s$, $\overline{\tau}_{ij}{}^s$ for the jth set of terminally admissible variations determined, in the sense of § 3, by the jth set of constants u_{hj}, in contradistinction to the notation $\gamma_j{}^s$, $\tau_{ij}{}^s$ for the set of end values of the jth set of differentially admissible variations. Then by means of (3.6), the column (4.8) may be written as

$$\begin{array}{lll} V_j & & 1 \text{ row} \\ \gamma_j{}^s - \overline{\gamma}_j{}^s & & 2 \text{ rows} \\ \overline{y}_i{}'^s(\gamma_j{}^s - \overline{\gamma}_j{}^s) + (\tau^s{}_{ij} - \overline{\tau}^s{}_{ij}). & & 2n \text{ rows} \end{array} \tag{4.8a}$$

We can simplify these columns by multiplying the second and third rows respectively by $\overline{y}_i{}'^s$ and substracting from the last $2n$ rows.

We thus obtain the necessary condition

$$
\begin{vmatrix}
V_2 \cdots\cdots\cdots\cdots\cdots\cdots V_{2n+3} \\
\gamma_1{}^1 - \bar{\gamma}_1{}^1 \cdots\cdots\cdots \gamma^1{}_{2n+3} - \bar{\gamma}^1{}_{2n+3} \\
\gamma_1{}^2 - \bar{\gamma}_1{}^2 \cdots\cdots\cdots \gamma^2{}_{2n+3} - \bar{\gamma}^2{}_{2n+3} \\
\eta_i{}^1{}_{11} - \bar{\eta}_i{}^1{}_{11} \cdots\cdots \eta_i{}^1{}_{1\,2n+3} - \bar{\eta}_i{}^1{}_{1\,2n+3} \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\eta_i{}^1{}_{n1} - \bar{\eta}_i{}^1{}_{n1} \cdots\cdots \eta_i{}^1{}_{n\,2n+3} - \bar{\eta}_i{}^1{}_{n\,2n+3} \\
\eta_i{}^2{}_{11} - \bar{\eta}_i{}^2{}_{11} \cdots\cdots \eta_i{}^2{}_{1\,2n+3} - \bar{\eta}_i{}^2{}_{1\,2n+3} \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\eta_i{}^2{}_{n1} - \bar{\eta}_i{}^2{}_{n1} \cdots\cdots \eta_i{}^2{}_{n\,2n+3} - \bar{\eta}_i{}^2{}_{n\,2n+3}
\end{vmatrix} = 0. \tag{4.9}
$$

Here $\gamma_j{}^s$, $\eta_{ij}{}^s$ are the end values of a set of differentially admissible variations, otherwise unrestricted, and $\bar{\gamma}_j{}^s$, $\bar{\eta}_{ij}{}^s$ is any set of terminally admissible variations.

Hence for any particular choice of $2n + 3$ sets of differentially admissible variations and of $2n + 3$ sets of constants u_{hj} $(j = 1, \cdots, 2n + 3)$, there must exist $2n + 3$ constants $(\lambda_0, b^s, k_i{}^s)$ not all zero such that the equation

$$\lambda_0 V(\gamma, \eta_i, u) + b^s(\gamma^s - \bar{\gamma}^s) + k_i{}^s(\eta_i{}^s - \bar{\eta}_i{}^s) = 0 \tag{4.10}$$

is satisfied for every set γ^s, $\eta_i(x)$ of the chosen sets of differentially admissible variations, and for every set u_h of the chosen sets of constants u_{hj}. Here $\bar{\gamma}^s$, $\bar{\eta}_i{}^s$ is the set of terminally admissible variations determined by u_h.

In fact, if we choose the $2n + 3$ constants $(\lambda_0, b^s, k_i{}^s)$ corresponding to the choice of $\gamma_j{}^s$, $\eta_{ij}(x)$, u_{hj} which gives the determinant (4.9) its highest rank ρ, then (4.10) must be satisfied for any differentially admissible variations and any set of numbers u_h whatever, $\bar{\gamma}^s$, $\bar{\eta}_i{}^s$ being the terminally admissible variations determined by u_h. For otherwise one of the columns of (4.9) could be replaced by another making the determinant have the rank $\rho + 1$.

(a) *We shall now obtain* (4.17) *and* (4.19) *as consequences of* (4.10). *We call attention to the fact that the derivation of* (4.17) *and* (4.19) *from* (4.10) *will hold even if E is not a minimizing arc, in fact if E is simply an arbitrary admissible arc for which* (4.10) *holds as stated.*

If we substitute in (4.10) the value of $\lambda_0 V(\gamma, \eta, u)$ given by (4.5), we obtain the equation

$$(4.11)\quad -\int_{a^1}^{a^2}\lambda_\tau\zeta_{\tau\tau}dx+(\lambda_0 f^s\gamma^s+\eta_{i^s}F_{y_{i'}{}^s})_1^2+\lambda_0\theta_h u_h + b^s(\gamma^s-\bar{\gamma}^s)+k_{i^s}(\eta_{i^s}-\bar{\eta}_{i^s})=0.$$

This equation must be satisfied for every choice of differentially admissible variations γ^s, $\eta_{i^s}(x)$; that is, for arbitrary γ^s, ζ_τ, and η_{i^s}. The values of $F_{y_{i'}{}^1}$ are also arbitrary. Furthermore, u_h is any set of r numbers, and $\bar{\gamma}^s$, $\bar{\eta}_{i^s}$ is the set of terminally admissible variations determined by u_h.

We can rewrite (4.11) as

$$(4.12)\quad -\int_{a^1}^{a^2}\lambda_\tau\zeta_{\tau\tau}dx+\gamma^s[(-1)^s\lambda_0 f^s+b^s]+\eta_{i^s}[(-1)^s F_{y_{i'}{}^s}+k_{i^s}] - b^s\bar{\gamma}^s-k_{i^s}\bar{\eta}_{i^s}+\lambda_0\theta_h u_h=0.$$

Since the value of $F_{y'_{i}{}^1}$ is arbitrary, we can choose

$$(4.13)\qquad F_{y_{i'}{}^1}=k_{i^1}$$

If we choose

$$\gamma^s=\eta_{i^2}=u_h=0,$$

thus determining $\bar{\gamma}^s$ and $\bar{\eta}_{i^s}$ as zero, we find that since ζ_τ is arbitrary, we must have the identities $\lambda_{m+1}\equiv\cdots\equiv\lambda_n\equiv 0$. Since γ^2, γ^1, and η_{i^2} are arbitrary, we find that

$$(4.14)\quad b^2=-\lambda_0 f^2=-F^2\qquad b^1=\lambda_0 f^1=F^1\qquad k_{i^2}=-F_{y_{i'}{}^2}.$$

Substituting these values for b^s, k_{i^s} and setting $\lambda_\tau=0$, (4.12) becomes

$$(4.15)\qquad (F^s\bar{\gamma}^s+F_{y_{i'}{}^s}\bar{\eta}_{i^s})_1^2+\lambda_0\theta_h u_h=0.$$

Recalling (3.6), we can write

$$\bar{\gamma}^s=dx^s\qquad \bar{\eta}_{i^s}=dy_{i^s}-\bar{y}_{i'}{}^s dx^s\qquad \theta_h u_h=d\theta$$

From (4.15) we now obtain the "transversality" condition* in the form of the following identity in the differentials $d\alpha_h$:

$$(4.16)\qquad [F^s dx^s+F_{y_{i'}{}^s}(dy_{i^s}-\bar{y}_{i'}{}^s dx^s)]_1^2+\lambda_0 d\theta=0$$

* The transversality condition given by Bliss in his lectures, loc. cit. p. 21, can be simply derived from this condition, but this condition, if derivable at all from the Bliss form, would clearly be so derivable only after complicated transformations. This is partly due to the presence of θ, and partly to the greater generality of the end conditions.

where dx^s, dy_i^s, and $d\theta$ are understood as evaluated for $(\alpha) = (0)$, and expressed in terms of the differentials $d\alpha_h$.

The multipliers λ_0, $\lambda_\beta(x)$ cannot all vanish at any one point of (a^1a^2). For if $\lambda_0 = 0$, the differential equations which the v's satisfy are homogeneous, and hence if the v's are zero for any one value of x, they are identically zero on the interval between corners containing that value of x. Thus they are identically zero on (a^1a^2). But this makes the constants λ_0, b^s, k_i^s all zero, as follows from (4.13) and (4.14), and so is impossible.

A set of multipliers λ_0, $\lambda_\beta(x)$, where λ_0 is a constant and $\lambda_1(x)$, $\cdots$, $\lambda_m(x)$ are functions of x which are continuous except possibly at the corners, and where λ_0, $\lambda_\beta(x)$ are not all zero at any one point of (a^1a^2), will be called an *admissible* set of multipliers.

The results just attained may be summarized in the following theorem:

THEOREM 2. *If E affords a minimum in the problem, there exists a set of constants c_1, $\cdots$, c_n and an admissible set of multipliers λ_0, $\lambda_1(x)$, $\cdots$, $\lambda_m(x)$ such that the equations*

$$F_{y_i'} = \int_{a^1}^{x} F_{y_i} dx + c_i \qquad A \tag{4.17}$$

are satisfied at every point of E, where

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_\beta \phi_\beta, \tag{4.18}$$

and such that the following identity

$$[(F^s - \bar{y}_i'^s F_{y_i'^s})dx^s + F_{y_i'^s} dy_i^s]_1^2 + \lambda_0 d\theta = 0 \qquad B \tag{4.19}$$

in $d\alpha_h$ is satisfied when dx^s, dy_i^s, and $d\theta$ are evaluated for $(\alpha) = (0)$ and expressed in terms of the differentials $d\alpha_h$.

An immediate consequence of this theorem is the so-called Euler-Lagrange Multiplier Rule:

On every sub-arc between corners of a minimizing arc the equations

$$\frac{d}{dx} F_{y_i'} = F_{y_i} \qquad \phi_\beta(x, y, y') = 0 \tag{4.20}$$

must be satisfied, where F is the function (4.18).

5. **Relative normalcy. A new criterion.** An admissible arc E shall be said to be *normal relative to conditions A and B of Theorem*

2, or, more shortly, *normal* (AB),* if it does not satisfy these conditions with any admissible set of multipliers in which $\lambda_0 = 0$.

A differentially admissible arc with an admissible set of multipliers is called an *extremaloid* if it satisfies equations (4.17) with some set of constants $c_1, \cdots, c_n$.

THEOREM 3. *For an extremaloid E which is normal (AB) and which satisfies A and B with an admissible set of multipliers, there always exist admissible multipliers in the form $\lambda_0 = 1, \lambda_\beta(x)$ with which E satisfies conditions A and B. In this form the multipliers are unique.*

We have multipliers in the form $\lambda_0, \lambda_\beta(x)$ where $\lambda_0 \neq 0$, and we can evidently divide them by λ_0 and obtain a set with $\lambda_0 = 1$. If we had two sets with $\lambda_0 = 1$, then the difference of the sets would be a set of multipliers with $\lambda_0 = 0$. This is impossible unless the two sets are identical, if the extremaloid is normal (AB). Thus the multipliers $\lambda_0 = 1, \lambda_\beta(x)$ are unique in this form.

Let us now consider the determinant

$$
(5.1) \qquad \begin{vmatrix}
\gamma_1{}^1 - \bar{\gamma}_1{}^1 \cdots \gamma^1{}_{2n+2} - \bar{\gamma}^1{}_{2n+2} \\
\gamma_1{}^2 - \bar{\gamma}_1{}^2 \cdots \gamma^2{}_{2n+2} - \bar{\gamma}^2{}_{2n+2} \\
\eta_i{}^1{}_{11} - \bar{\eta}_i{}^1{}_{11} \cdots \eta_i{}^1{}_{1\,2n+2} - \bar{\eta}_i{}^1{}_{1\,2n+2} \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
\eta_i{}^1{}_{n1} - \bar{\eta}_i{}^1{}_{n1} \cdots \eta_i{}^1{}_{n\,2n+2} - \bar{\eta}_i{}^1{}_{n\,2n+2} \\
\eta_i{}^2{}_{11} - \bar{\eta}_i{}^2{}_{11} \cdots \eta_i{}^2{}_{1\,2n+2} - \bar{\eta}_i{}^2{}_{1\,2n+2} \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
\eta_i{}^2{}_{n1} - \bar{\eta}_i{}^2{}_{n1} \cdots \eta_i{}^2{}_{n\,2n+2} - \bar{\eta}_i{}^2{}_{n\,2n+2}
\end{vmatrix}
$$

formed from $2n + 2$ sets of differentially admissible variations

$$\gamma_j{}^s, \eta_{ij}(x) \qquad (i = 1, \cdots, n) \quad (j = 1, \cdots, 2n + 2; s = 1, 2)$$

and $2n + 2$ sets of terminally admissable variations

$$\bar{\gamma}_j{}^s, \bar{\eta}^s{}_{ij} \qquad (i = 1, \cdots, n; j = 1, \cdots, 2n + 2; s = 1, 2)$$

* We say *normal* (AB) rather than simply *normal* for the following reason. Morse has shown that Theorem 2 holds with the following condition C added at each corner $x = c$,

$$[F - y_i' F_{y_i'}]_{c+}^{c-} = 0 \qquad C$$

and that moreover the multipliers in this new theorem do not always agree with those in the old theorem. He has shown that an admissible arc may be normal (ABC)—that is, relative to conditions A, B, and C—while anormal (AB).

and the determinant

$$(5.2)\qquad \begin{vmatrix} \eta_i^1{}_{11} - \bar{\eta}_i^1{}_{11} \cdots \eta_i^1{}_{1\,2n} - \bar{\eta}_i^1{}_{1\,2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \eta_i^1{}_{n1} - \bar{\eta}_i^1{}_{n1} \cdots \eta_i^1{}_{n\,2n} - \bar{\eta}_i^1{}_{n\,2n} \\ \eta_i^2{}_{11} - \bar{\eta}_i^2{}_{11} \cdots \eta_i^2{}_{1\,2n} - \bar{\eta}_i^2{}_{1\,2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \eta_i^2{}_{n1} - \bar{\eta}_i^2{}_{n1} \cdots \eta_i^2{}_{n\,2n} - \bar{\eta}_i^2{}_{n\,2n} \end{vmatrix}$$

formed from $2n$ sets of differentially admissible variations

$$\gamma_k{}^s,\ \eta_{ik}(x) \qquad (i = 2, \cdots, n;\ k = 1, \cdots, 2n;\ s = 1, 2)$$

and $2n$ sets of terminally admissible variations

$$\bar{\gamma}_k{}^s,\ \bar{\eta}_i{}^s{}_{ik} \qquad (i = 1, \cdots, n;\ k = 1, \cdots, 2n;\ s = 1, 2).$$

We will now prove the following lemma:

LEMMA. *A necessary and sufficient condition that there exist a determinant of form* (5.1) *different from zero is that there exist a determinant of form* (5.2) *different from zero.*

The sufficiency of this condition is proved by showing that if there exists a determinant of form (5.2) different from zero, then we can construct a determinant (5.1) different from zero.

In fact, if a_{pq} is the general element of the determinant (5.2), by choosing the arbitrary constants $\gamma_j{}^s$ of (5.1) suitably we obtain a determinant of the form

$$(5.3)\qquad \left|\begin{array}{c|cc} 0\cdots\cdots 0 & 0 & a \\ 0\cdots\cdots 0 & b & 0 \\ \hline & & \\ a_{pq} & & \\ & & \end{array}\right| \qquad a \neq 0,\ b \neq 0$$

different from zero.

Conversely, if there exists a determinant of form (5.1) not zero, then determinants of form (5.2) obviously cannot all be zero. Thus the condition is also necessary.

The following theorem relates to the differential equations (3.3) alone, and is independent of whether E is or is not a minimizing arc.

THEOREM 4. *A necessary and sufficient condition that an admissible arc E be normal (AB) is that there exist for it a determinant of form* (5.1) *different from zero.*

If all determinants of form (5.1) vanished, there would clearly be a set of constants $\lambda_0 = 0$, b^s, k_i^s satisfying (4.10) for any sets of differentially and terminally admissible variations. But as pointed out in statement (a) of § 4, conditions A and B on E would then follow as consequences of (4.10), λ_0 being zero in this case. Thus E is anormal (AB) if all determinants of form (5.1) vanish, and the condition is proved necessary.

Now suppose that the arc E were anormal (AB). We shall show that there exists no determinant of form (5.1) different from zero.

Suppose that there were a determinant of form (5.1) *different from zero.* Then there would also be a determinant of form (5.2) different from zero. By the definition of normalcy, conditions A and B would be satisfied by E with a set of multipliers in which $\lambda_0 = 0$. The function F in (4.17) and (4.19) would then have the form

$$F = \lambda_\beta \phi_\beta \tag{5.4}$$

Every set of differentially admissible variations would satisfy the equations

$$0 = \int_{a^2}^{a^3} \lambda_\beta \Phi_\beta dx = \int_{a^1}^{a^2} (F_{y_i} \eta_i + F_{y_i'} \eta_i') dx. \tag{5.5}$$

Upon integration by parts, and application of (4.17), we have

$$(F_{y_i'^s} \eta_i{}^s)_1^2 = 0. \tag{5.6}$$

Since we are assuming $\lambda_0 = 0$, and since $\phi_\beta = 0$ along E, we have $F^s = 0$. We then see from the transversality condition (4.16) that

$$(F_{y_i'^s} \bar{\eta}_i{}^s)_1^2 = 0 \tag{5.7}$$

and this would hold for any terminally admissible variations $\bar{\eta}_i{}^s$. Combining (5.6) and (5.7), we would obtain the equation

$$[F_{y_i'^s} (\eta_i{}^s - \bar{\eta}_i{}^s)]_1^2 = 0 \tag{5.8}$$

But we have a determinant of form (5.2) different from zero, so that from (5.8) we deduce the equations

$$F_{y_i'^s} = 0 \qquad (s = 1, 2; i = 1, \cdots, n). \tag{5.9}$$

But equations (5.9) would imply that $\lambda_\beta(a^s) = 0$, which is contrary to our hypothesis that $\lambda_0, \lambda_\beta(x)$ form an admissible set of multipliers. Thus there cannot be a determinant of form (5.1) different from zero if E is anormal (AB). This proves the sufficiency of the condition.

COROLLARY. *A necessary and sufficient condition that an admissible arc E be normal (AB) is that there exist for it a determinant of form* (5.2) *different from zero.*

THEOREM 5. *If a set of differentially admissible variations $\gamma^s, \eta_i(x)$ for an admissible arc E which is normal (AB) satisfies the terminal equations of variation with a set (u), then there exists a one-parameter family of admissible arcs of the form* (3.7) *containing E for the parameter value $t = 0$ and having $\gamma^s, \eta_i(x)$, u_h as the variations of x^s, y_i, and α_h.*

According to the preceding corollary we have a determinant of form (5.1) different from zero, with its $2n + 2$ sets of differentially admissible variations

$$\gamma_j{}^s, \eta^s{}_{ij} \qquad (j = 1, \cdots, 2n + 2)$$

and its $2n + 2$ sets of terminally admissible variations

$$\bar{\gamma}_j{}^s, \bar{\eta}^s{}_{ij} \qquad (j = 1, \cdots, 2n + 2) \tag{5.10}$$

Let

$$u_{hj} \qquad (h = 1, \cdots, r; j = 1, \cdots, 2n + 2)$$

be constants of which the jth set determines the jth set in (5.10), in the sense of § 3.

There exists a $(2n + 3)$-parameter family of differentially admissible arcs

$$y_i = y_i(x, t, e) \qquad x_0{}^1(t, e) \leqq x \leqq x_0{}^2(t, e) \tag{5.11}$$

with parameters t and $(e) = (e_1, \cdots, e_{2n+2})$, containing E for the parameter values $(t, e) = (0, 0)$ and having the sets $\gamma^s, \eta_i(x)$ and $\gamma_j{}^s, \eta_{ij}(x)$ as its variations along E with respect to t and e_j respectively.

Let us take the parameters α_h of (2.4) as follows:

$$\alpha_h = tu_h + e_j u_{hj} \qquad (h = 1, \cdots, r; j = 1, \cdots, 2n + 2).$$

Then the equations

$$\begin{aligned} x_0{}^s(t, e) - x^s(tu_1 + e_j u_{1j}, \cdots, tu_r + e_j u_{rj}) &= 0 \\ y_i[x_0{}^s(t, e), t, e] - y_i{}^s(tu_1 + e_j u_{1j}, \cdots, tu_r + e_j u_{rj}) &= 0, \end{aligned} \tag{5.12}$$

which are conditions for terminal admissibility, have the initial solution $(t, e) = (0, 0)$, at which the jacobian with respect to (e) is the non-vanishing determinant (5.1), as is seen by comparison with (4.7), (4.8) and (4.9). Hence equations (5.12) have solutions

$$e_j = e_j(t) \qquad (j = 1, \cdots, 2n + 2) \tag{5.13}$$

with initial values $e_j\,(0) = 0$. The one-parameter family

$$y_i = y_i[x, t, e(t)] \qquad x_0{}^1[t, e(t)] \leqq x \leqq x_0{}^2[t, e(t)] \tag{5.14}$$

consists of admissible arcs, and contains E for $t = 0$.

Let us compute the derivatives $e_j'(0)$. From the choice of γ^s, $\tau_{ii}(x)$ and u_h we see that the partial derivatives of the left hand sides of (5.12) with respect to t are zero for $t = 0$, and therefore $e_j'(0) = 0$. We see then that the variations of the family (5.14) with respect to t are γ^s, $\tau_{ii}(x)$, u_h, as was to be proved.

6. **Geometric criteria for normalcy (AB).** Let $\gamma^s, \tau_{ii}{}^s$, be the end-values of a set of differentially admissible variations. Then we will consider the $2n + 2$ numbers $\gamma^s, \tau_{ii}{}^s$, as the coördinates of a point in $(2n + 2)$-dimensional space. Such a point will be called a *differentially admissible point.* Similarly, consider any set of terminally admissible variations $\bar{\gamma}^s, \bar{\tau}_{ii}{}^s$ as the coördinates of a point in $(2n + 2)$-space. Such a point will be called a *terminally admissible point.* The following theorem furnishes a useful geometrical interpretation of normalcy (AB).

THEOREM 6. *A necessary and sufficient condition that an admissible arc E be normal (AB) is that the family of $(2n + 2)$-dimensional points linearly dependent on the differentially admissible points and the terminally admissible points fill out the whole of $(2n + 2)$-space.*

For if an admissible arc is normal (AB) there exists a determinant of form (5.1) different from zero. The columns of this determinant are $2n + 2$ linearly independent points of the family; hence the family fills out $(2n + 2)$-space.

Now suppose that the family fills out $(2n + 2)$-space. Then there exist $2n + 2$ linearly independent points in the family. But any point of the family may be expressed as the difference of a differentially admissible point and a terminally admissible point because of the linear homogeneity of the differential and terminal equations of variation. Therefore there exists a determinant of form (5.1) not zero, and the arc E is normal AB.

In a similar manner the following theorem can be proved:

THEOREM 7. *A necessary and sufficient condition that an admissible arc be normal (AB) is that the family of 2n-dimensional points linearly dependent on the points in 2n-space with coördinates $\tau_{i_i}{}^s$ and the points in 2n-space with coördinates $\bar{\tau}_{i_i}{}^s$ fill out the whole of 2n-space. Here the set $\tau_{i_i}{}^s$ consists of the end values of differentially admissible variations τ_{i_i}, and the set $\bar{\tau}_{i_i}{}^s$ consists of terminally admissible variations.*

The following test for normalcy (*AB*) is useful:

THEOREM 8. *Let ρ be the highest rank attained by any determinant consisting of $2n+2$ columns of differentially admissible points in $(2n+2)$-space. The number of parameters (α) in the terminal conditions* (2.4) *cannot be less than $2n+2-\rho$ without E being anormal (AB). One can, however, always choose terminal conditions of the form* (2.4) *with no more than $2n+2-\rho$ parameters such that E is normal (AB).*

The differentially admissible points fill out a ρ-plane, and the terminally admissible points must fill out at least the $(2n+2-\rho)$-plane orthogonal to the ρ-plane at the origin if E is to be normal (AB). This latter plane is certainly not filled out if the number of parameters is less than $2n+2-\rho$. This proves the first part of the theorem.

Now the $(2n+2-\rho)$-plane which must be filled out by the differentially admissible points if the arc E is to be normal (AB) may be represented in the following form, using $(\bar{\gamma}^s, \bar{\tau}_{i_i}{}^s)$ as coördinates in $(2n+2)$-space:

$$\begin{aligned} \bar{\gamma}^s &= c_h{}^s u_h \\ \bar{\tau}_{i_i}{}^s &= c_{ih}{}^s u_h. \end{aligned} \tag{6.1}$$

The matrix $\left\| \begin{matrix} c_h{}^s \\ c_{ih}{}^s \end{matrix} \right\|$ is to be of rank $2n+2-\rho$. A possible choice of terminal conditions which will produce (6.1) as the terminally admissible points is the following: (cf. (3.6))

$$\begin{aligned} x^s - a^s &= c_h{}^s \alpha_h \qquad (h = 1, \cdots, 2n+2-\rho) \\ y_i{}^s - b_i{}^s &= (\bar{y}_i{}'{}^s c_h{}^s + c^s{}_{ih})\alpha_h \end{aligned} \tag{6.2}$$

where $(a^s, b_i{}^s)$ are the coördinates of the end points of E.

A theorem similar to the last one can be obtained by replacing $2n+2-\rho$ by $2n-\sigma$, where σ is the highest rank attained by any determinant of the form

$$\begin{vmatrix} \eta^1{}_{11} \cdots \eta^1{}_{1\,2n} \\ \cdots\cdots\cdots\cdots \\ \eta^1{}_{n1} \cdots \eta^1{}_{n\,2n} \\ \eta^2{}_{11} \cdots \eta^2{}_{1\,2n} \\ \cdots\cdots\cdots\cdots \\ \eta^2{}_{n1} \cdots \eta^2{}_{n\,2n} \end{vmatrix} . \tag{6.3}$$

A differentially admissible arc shall be said to be *normal* (A) if it does not satisfy condition A of Theorem 2 with any admissible set of multipliers in which $\lambda_0 = 0$.

THEOREM 10.* *If a differentially admissible arc E is normal (A), there exists a family of differentially admissible arcs joining the points (x^1, y^1) and (x^2, y^2) neighboring the end points of E and representable in the form*

$$y_i = Y_i(x, x^1, y^1, x^2, y^2) \tag{6.4}$$

where the functions Y_i are of class C'.

Since E is normal (A), it is normal (AB) if we choose the end-conditions so as to make the end points coincide with those of E. Hence, according to the corollary to Theorem 4, there exists a determinant of end values of differentially admissible variations of the form

$$\left|\eta^s{}_{ik}\right| \quad (i = 1, \cdots, n;\ k = 1, \cdots, 2n) \tag{6.5}$$

different from zero.

Now let us use the corollary of § 3 to set up a 2n-parameter family of differentially admissible arcs

$$y_i = y_i(x, e) \qquad (e) = (e_1, \cdots, e_{2n}) \tag{6.6}$$

whose variations with respect to e_k are precisely the variations in the k-th column of (6.5) and which contains the arc E for $(e) = (0)$. We then seek to satisfy the conditions

$$y_i{}^s = y_i(x^s, e). \tag{6.7}$$

These conditions are satisfied by the end points of E and $(e) = (0)$, and the jacobian of the right hand members with respect to (e) is the determinant (6.5), different from zero. We can accordingly solve

* cf. Bliss, Lectures, loc. cit. p. 50.

(6.7) for (e), substitute the result in (6.6), and obtain the desired family.

References

[1] Morse. For an abstract see "The Problems of Lagrange and Mayer under General End Conditions," Proceedings of the National Academy of Sciences, vol. 16 (1930), pp. 229–233.

[2] Bolza. "Uber den Anormalen Fall beim Lagrangeschen und Mayerschen Problem mit gemischten Bedingungen und Variabeln Endpunkten," Mathematische Annalen, vol. 74 (1913), pp. 430–446.

[3] Bliss. "The Problem of Lagrange in the Calculus of Variations," lectures given at the University of Chicago, summer quarter, 1925. These lectures were prepared by O. E. Brown, Northwestern University, Evanston, Illinois.

Further references follow:

Radon. "Zum Problem von Lagrange," Abhandlungen aus dem Mathematischen Seminar, Hamburg, 1928.

Bliss. "The Problem of Mayer with Variable End Points," Transactions of the American Mathematical Society," vol. 19 (1918), p. 312.

Bolza. "Vorlesungen über Variationsrechnung."

Bliss. "The Problem of Lagrange in the Calculus of Variations," American Journal of Mathematics, vol. 52 (1930), pp. 673–744.